

Titre: Estudio de diferentes parámetros para la detección de esclerosis lateral amiotrófica a partir del movimiento articulario
Title: lateral amiotrófica a partir del movimiento articulario

Auteurs: Cristina Carmona, Pedro Gómez, Miguel A. Ferrer, Réjean Plamondon, & Ana Londral
Authors: Plamondon, & Ana Londral

Date: 2017

Type: Article de revue / Article

Référence: Carmona, C., Gómez, P., Ferrer, M. A., Plamondon, R., & Londral, A. (2017). Estudio de diferentes parámetros para la detección de esclerosis lateral amiotrófica a partir del movimiento articulario. Loquens, 4 (1), 1-5.
Citation: <https://doi.org/10.3989/loquens.2017.038>

Document en libre accès dans PolyPublie

Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/4815/>
PolyPublie URL:

Version: Version officielle de l'éditeur / Published version
Révisé par les pairs / Refereed

Conditions d'utilisation: CC BY
Terms of Use:

Document publié chez l'éditeur officiel

Document issued by the official publisher

Titre de la revue: Loquens (vol. 4, no. 1)
Journal Title:

Maison d'édition: Editorial CSIC
Publisher:

URL officiel: <https://doi.org/10.3989/loquens.2017.038>
Official URL:

Mention légale:
Legal notice:

Estudio de diferentes parámetros para la detección de esclerosis lateral amiotrófica a partir del movimiento articulatorio

Cristina Carmona¹, Pedro Gómez², Miguel A. Ferrer¹, Réjean Plamondon³, Ana Londral⁴

¹ Instituto Universitario para el Desarrollo Tecnológico y la Innovación en Comunicaciones, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, España

² Center for Biomedical Technology, Universidad Politécnica de Madrid, Campus de Montegancedo, s/n, 28223 Pozuelo de Alarcón, Madrid, España

³ Laboratoire Scribens, Département de Génie Électrique, École Polytechnique de Montréal, Canadá

⁴ Instituto de Medicina Molecular, Universidade de Lisboa, Lisboa, Portugal

ccarmona@idetec.eu ORCID: <http://orcid.org/0000-0002-4441-6652>

pedro@fi.upm.es ORCID: <http://orcid.org/0000-0003-3283-378X>

mferrer@idetec.eu ORCID: <http://orcid.org/0000-0002-2924-1225>

rejean.plamondon@polymtl.ca ORCID: <http://orcid.org/0000-0002-4903-7539>

analondral@gmail.com ORCID: <http://orcid.org/0000-0002-8002-6790>

Enviado: 29/02/2016. Aceptado: 30/12/2016. Publicado online: 18/12/2017

Citation / Cómo citar este artículo: Carmona, C., Gómez, P., Ferrer, M. A., Plamondon, R. y Londra, A. (2017). Estudio de diferentes parámetros para la detección de esclerosis lateral amiotrófica a partir del movimiento articulatorio. *Loquens*, 4(1), e038. doi: <http://dx.doi.org/10.3989/loquens.2017.038>

RESUMEN: La esclerosis lateral amiotrófica (ELA) es una enfermedad degenerativa de tipo neuromuscular, uno de cuyos primeros síntomas es la dificultad para hablar. Recientemente se ha presentado un nuevo método para determinar el movimiento articulatorio a partir de la estimación de los formantes. En este artículo se presentan dos parámetros obtenidos mediante el modelado del movimiento articulatorio y se evalúan con respecto al grado de la enfermedad.

Palabras clave: ELA; Sigma-lognormal; formantes; movimiento articulatorio.

ABSTRACT: *Study of several parameters for the detection of amyotrophic lateral sclerosis from articulatory movement.* – Amyotrophic lateral sclerosis (ALS) is a degenerative neuromuscular disease, one of its early symptoms being a progressive difficulty to speak (ALS dysarthria). To improve its diagnosis and monitoring, a new method based on articulatory movement estimation has been developed. As a result, two articulatory movement parameters are presented as well as their relationship with the illness grade.

Keywords: ALS; Sigma-lognormal; formant; kinematic; articulation.

1. INTRODUCCIÓN

La esclerosis lateral amiotrófica (ELA) es una enfermedad neurodegenerativa en la que el funcionamiento de las neuronas motoras va decreciendo gradualmente hasta que mueren, lo que produce una disminución paulatina en la capacidad de movimiento del paciente (Gómez-Vilda *et al.*, 2015). En la voz se manifiesta al principio por una dificultad progresiva para articular, hasta que los afectados ya no pueden comunicarse. El pronóstico de esta enfermedad es grave, pero si se detecta a tiempo se puede ralentizar el progreso de la enfermedad mediante medica-

ción, aumentando la esperanza y la calidad de vida del paciente.

Cuando esta enfermedad afecta en primer lugar a la fonación, uno de los principales problemas para diagnosticarla es que puede ser confundida con otro tipo de patología de los pliegues vocales. Esto es debido a que los primeros síntomas no resultan evidentes para los actuales métodos de diagnóstico de la voz (Delgado, 2016) y ello hace que demore el diagnóstico correcto y por tanto se retrase el tratamiento adecuado para el paciente.

Para facilitar la discriminación entre las patologías que afectan a los pliegues vocales por el cambio físico en

sus estructuras y las producidas por una enfermedad neurodegenerativa, sería recomendable emplear técnicas que permitan analizar el movimiento muscular del sistema articulatorio y fonatorio. Para este fin, se ha desarrollado en este trabajo un método para obtener la cinemática del movimiento de la lengua y de la mandíbula (Carmona-Duarte, Alonso *et al.*, 2016; Carmona-Duarte *et al.*, 2017) a partir de los formantes del habla. Este procedimiento consiste en averiguar la velocidad y la posición del sistema articulatorio mediante los dos primeros formantes (f_1 y f_2). A partir de estos últimos se puede inferir la velocidad de un punto de referencia en el sistema biomecánico lengua-mandíbula. La velocidad resultante puede parametrizarse por el modelo de movimientos rápidos Sigma-Lognormal (Plamondon, 1995), que ha sido empleado con éxito en el modelado de la escritura (Gomez-Barrero *et al.*, 2013; O'Reilly y Plamondon, 2012; Plamondon, Djoua y Mathieu, 2013; Plamondon, O'Reilly y Ouellet-Plamondon, 2014; Plamondon, O'Reilly, Rémi y Duval, 2013). Con este modelo se evalúan en este trabajo dos parámetros que indican la amplitud y el ritmo del movimiento cinemático del sistema lengua-mandíbula, lo que permite estimar el funcionamiento de las neuronas motoras del sistema cráneo-facial.

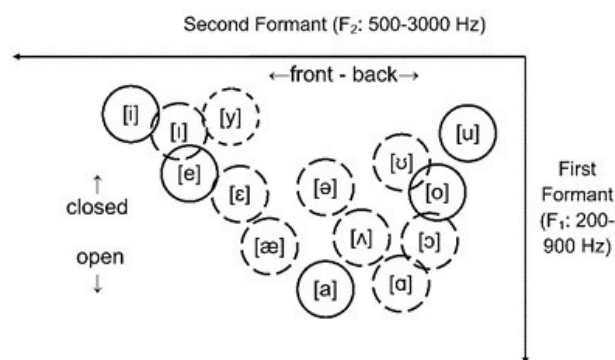
En el apartado 2 de este artículo, se presenta el método que se utiliza para obtener a partir de los formantes la señal de velocidad, y esta se descompone en órdenes motoras simples. En el apartado 3 se desarrolla la metodología y los parámetros que se utilizan para evaluar la voz. A continuación, en los apartados 4 y 5, se presentan los resultados obtenidos y las conclusiones, respectivamente.

2. ESTIMACIÓN DE LA CINEMÁTICA ARTICULATORIA Y APLICACIÓN DEL MODELO SIGMA-LOGNORMAL

Es bien conocido que las cavidades resonantes son modificadas por los órganos articulatorios, que crean diferentes sonidos dependiendo de la disposición que presentan. Si se considera que la suma de todos los movimientos musculares está sincronizada para fijar una determinada posición de referencia, se puede asumir que a partir de ella se pueden formar diferentes cavidades resonantes y, por tanto, producir un sonido vocálico diferente. Cada uno de estos sonidos se representa por una combinación diferente de las frecuencias correspondientes al primer y al segundo formante. Cada formante representa una concentración de energía en una determinada frecuencia producida por la cavidad de resonancia fijada. Así se puede relacionar esta posición de equilibrio con la fonación de las diferentes vocales (Peterson y Barney, 1952). Resulta así el familiar triángulo vocálico que se muestra en la Figura 1.

Recientes estudios (Carmona-Duarte, Alonso *et al.*, 2016; Carmona-Duarte, Plamondon *et al.*, 2016; Carmona-Duarte *et al.*, 2017; Gómez-Vilda *et al.*, 2015) relacionan la evolución de los dos primeros formantes, es decir, el in-

Figura 1: Representación del triángulo vocálico adaptado al castellano (círculos continuos) y al inglés americano (círculos discontinuos).



cremento de frecuencia del primer y del segundo formante (ΔF_1 y ΔF_2), con la evolución de la posición del centro de referencia citado mediante la siguiente expresión:

$$(1) \quad \begin{aligned} \Delta y &= K_1 \Delta F_1 \\ \Delta x &= K_2 \Delta F_2 \end{aligned}$$

De esta forma se podría decir que, si se pronuncia la secuencia /aiua/, se estaría dibujando un triángulo en el plano $\{f_2, f_1\}$. Si el sujeto repite el diptongo /ai/, se puede verificar cómo se dibuja un camino de la posición de /a/ a la de /i/ (Figura 2), y similarmente en los diptongos /au/ e /iu/.

Una vez transformados los formantes a unas posiciones en el espacio es posible obtener una estimación de la velocidad a la que se mueve el punto de referencia (variable cinemática). Ello se explica por la expresión:

$$(2) \quad \begin{aligned} |\vec{v}(t)| &= \frac{\sqrt{\Delta x(t)^2 + \Delta y(t)^2}}{\Delta t} = \\ &= \frac{\sqrt{(K_2 \Delta F_2)^2 + (K_1 \Delta F_1)^2}}{\Delta t} \end{aligned}$$

En la Figura 3 se puede observar cómo sería el perfil de velocidad para la producción repetitiva del diptongo /ai/.

Una vez se haya obtenido el perfil de velocidad, este puede explicarse mediante el modelo Sigma-Lognormal (Plamondon y Djoua, 2006), que establece que los movimientos rápidos pueden ser modelados por la suma de movimientos más simples, y que por tanto el perfil de velocidad se puede descomponer en patrones Sigma-Lognormales, que, al sumarse, forman el perfil de velocidad en módulo y en los trazos simples que componen el movimiento complejo. Por consiguiente, cada uno de los perfiles de velocidad ($\vec{v}_j(t)$) para el movimiento j^o del sistema muscular linguo-mandibular estará descrito por la función lognormal ($\Lambda(t)$) (O'Reilly y Plamondon, 2009):

Figura 2: Transformación de los formantes a la posición del punto de referencia para los diptongos /au/, /iu/ y /ai/.

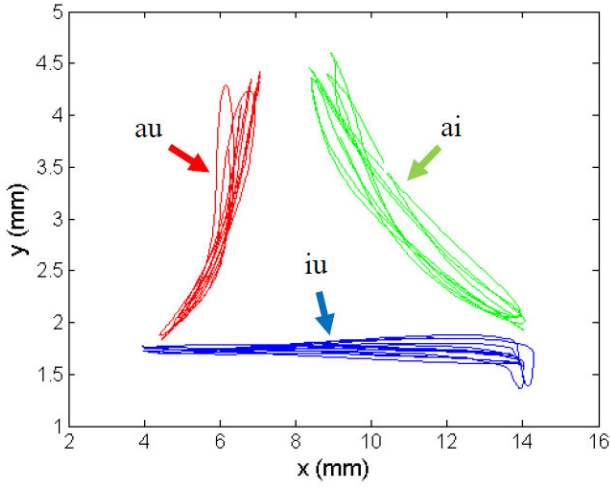
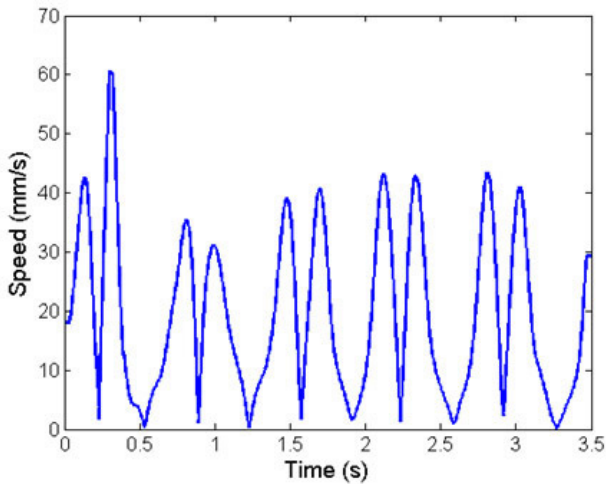


Figura 3: Perfil de velocidad para la producción repetitiva del diptongo /ai/.



$$(3) \quad \begin{aligned} |\vec{v}_j(t)| &= D\Lambda(t; t_0, \mu, \sigma) = \\ &= \frac{D}{\sqrt{2\pi}(t - t_0)} e^{-\frac{(\ln(t-t_0) - \mu)^2}{2\sigma^2}} \end{aligned}$$

donde t_0 es el instante de suceso de cada lognormal, D es el factor de escala, y μ y σ , los parámetros de media y desviación estándar, respectivamente.

El vector de velocidad se obtiene sumando cada una de las M distribuciones lognormales resultantes que se han aproximado previamente a cada uno de los picos de velocidad:

$$(4) \quad v_n(t) = \sum_{i=1}^M |\vec{v}_j(t)|$$

La estimación reconstruida por el modelo Sigma-Log-normal se compara con la estimación de velocidad calculada originalmente sobre el punto de referencia mediante la relación de señal a ruido (SNR), que viene definida en (O'Reilly y Plamondon, 2012):

$$(5) \quad SNR = 10 \log \left[\frac{\int_{\zeta=t_s}^{t_n} v(\zeta)^2 d\zeta}{\int_{\zeta=t_s}^{t_n} [v_n(\zeta) - v(\zeta)]^2 d\zeta} \right]$$

En el caso de que la SNR sea superior a 20 dB se puede considerar que la aproximación es suficientemente precisa, tal como se observa en la Figura 4 y se recoge en Carmona-Duarte, Alonso *et al.* (2016).

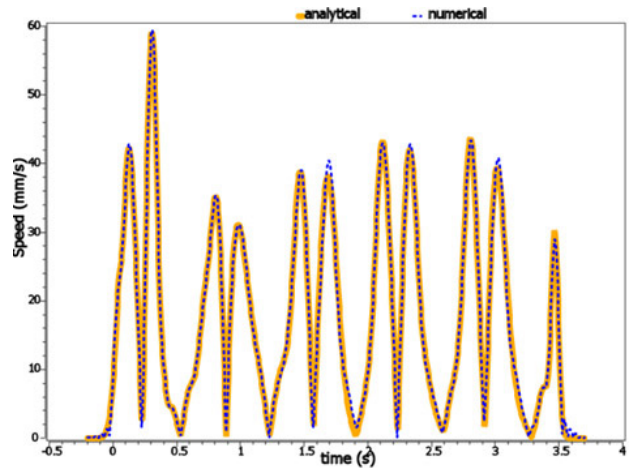
A partir de de la descomposición de la variable cinemática del punto de referencia en distribuciones lognormales y con la metodología que se expone en el apartado 3, se obtienen unos parámetros robustos para el diagnóstico de ELA, tal como se expone en la sección de resultados.

3. METODOLOGÍA

Para seleccionar el tipo de fonación más recomendado y para obtener los parámetros de ELA, se ha partido de la hipótesis siguiente: si el sujeto padece ELA de modo que afecte al movimiento del punto de referencia del sistema biomecánico motor participante en la articulación linguo-mandibular, sus movimientos empezarán a ser más lentos que en los sujetos normales y, por tanto, la distribución lognormal de cada movimiento simple tendrá menos amplitud y más dispersión que en los sujetos normales.

En los siguientes subapartados se detallan tanto el método de adquisición de datos y características de los sujetos participantes en el experimento de validación de la

Figura 4: Perfil de velocidad para la fonación repetitiva del diptongo /ai/ (línea continua) y su reconstrucción (línea discontinua).



hipótesis como los parámetros utilizados teniéndola en cuenta.

3.1. Sujetos y toma de datos

Para la comparación de los sujetos sanos con los patológicos se han tomados datos de habla continua, en la cual la persona lee de manera natural una frase. La toma de datos se ha llevado a cabo en portugués, a partir de la siguiente oración: “/tudo vale a pena quando a alma não é pequena/”. Las pruebas se han llevado a cabo tomando dos locutoras normativas (controles sanos) de diferentes edades, con 36 y 63 años respectivamente (CF36 y CF63), para compensar el sesgo que pueda introducir el envejecimiento natural, y dos pacientes femeninas con ELA bulbar en grado 1 y 2 (PF3 y PF2, respectivamente). Hay que tener en cuenta que resulta sumamente difícil encontrar pacientes en estadios incipientes de la enfermedad, lo que complica la adquisición de una base de datos de mayor tamaño.

Los datos fueron grabados en consulta médica con una frecuencia de muestreo de 16 kHz y 16 bits.

Una vez adquiridas las muestras, los formantes se obtuvieron con una resolución en frecuencia de 15 Hz cada 1 ms, aplicando un filtrado inverso adaptativo de orden 8 sobre la señal remuestreada a 8 kHz.

3.2. Parámetros

Para estimar los parámetros del modelo se partió de Carmona-Duarte, Plamondon *et al.* (2016), donde se observó que el parámetro más significativo estadísticamente en el caso del ELA es la diferencia media $\Delta\sigma$, siendo σ la desviación estándar definida en la expresión (3). Este parámetro (p_1) se estima según la expresión:

$$(6) \quad p_1 = \Delta\sigma = \frac{\sum_{i=1}^M |\sigma_i - \sigma_{i-1}|}{M}$$

En este trabajo, además, se ha propuesto otro parámetro (p_2) relacionado con $\Delta\sigma$, teniendo en cuenta la dispersión de los valores estimados:

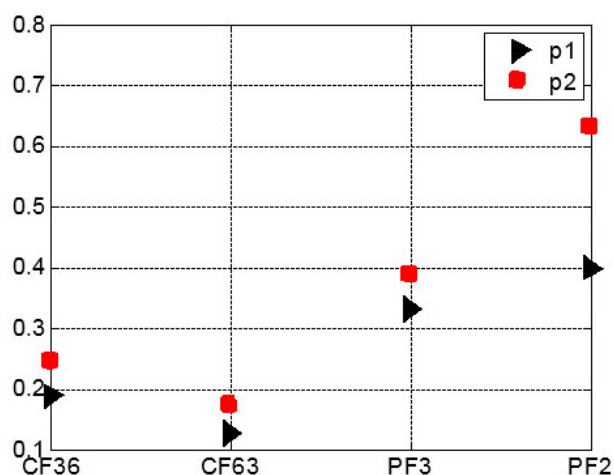
$$(7) \quad p_2 = Q_3(\Delta\sigma) - Q_1(\Delta\sigma)$$

siendo $Q_1(\Delta\sigma)$ y $Q_3(\Delta\sigma)$ los cuartiles primero y tercero de la distribución de σ , lo que se conoce como rango intercuartil (H-spread).

4. RESULTADOS

Para llegar a los resultados se procesó primeramente las grabaciones de sonido de los dos sujetos sanos y de

Figura 5: Resultados obtenidos con dos métodos diferentes (p_1 y p_2) para dos sujetos normales (C36 y CF63) y dos sujetos con ELA (PF3 y PF2).



los dos con ELA. Una vez obtenidos los formantes, se pasó a realizar el proceso expuesto en el apartado 2. Tras estimar el parámetro σ de cada distribución lognormal, se calcularon los parámetros p_1 y p_2 . En la Figura 5 se muestran los resultados de las estimaciones con el método propuesto.

En la Figura 5 se puede observar cómo para el control de mayor edad las distribuciones lognormales se hacen más estrechas (menos dispersas), mientras que en el caso de los pacientes con ELA (PF3 y PF2) se hacen mucho más anchas. También se aprecia cómo el parámetro p_2 da mayor diferencia entre los grados de la enfermedad con respecto al parámetro p_1 .

En la Tabla 1 se puede ver que el parámetro p_2 mejora la diferencia entre las clases sanas y con ELA de forma significativa.

Tabla 1: p value para cada parámetro entre sujetos sanos y con ELA.

Parámetro	p value
p_1	0.045
p_2	0.007

Según se observó en Carmona-Duarte, Plamondon *et al.* (2016), también se puede afirmar que el parámetro p_1 no presentaba ninguna diferencia estadística con respecto a los sujetos que tenían algún tipo de patología laríngea no relacionada con una enfermedad de origen neuromotor.

5. CONCLUSIONES

En este artículo se han presentado nuevos parámetros que pueden servir de ayuda en el diagnóstico de ELA. Se ha partido de una nueva metodología para poder analizar el movimiento muscular del sistema articulatorio linguo-

mandibular a partir de los formantes obtenidos de la señal de habla. El resultado se utilizó para la evaluación de sujetos con ELA y en la valoración de diversos grados de progreso de la enfermedad. La principal ventaja de estos parámetros es que no son sensibles a otros tipos de patologías, como pueden ser las de tipo laríngeo, ni al envejecimiento natural de la persona.

AGRADECIMIENTOS

Este estudio ha sido financiado por los proyectos del Gobierno español TEC2016-77791-C4-1-R y TEC2016-77791-C4-4-R (Plan Nacional de I+D, Ministerio de Economía y Competitividad), los fondos FEDER, un contrato postdoctoral de la Universidad de Las Palmas de Gran Canaria y el proyecto NSERC-Canada Grant RGPIN-2015-06409.

REFERENCIAS

- Carmona-Duarte, C., Alonso, J. B., Díaz, M., Ferrer, M. A., Gómez-Vilda, P., y Plamondon, R. (2016). Kinematic modeling of diphthong articulation. En A. Esposito, M. Faundez-Zanuy, A. M. Esposito, G. Cordasco, T. Drugman, J. Solé-Casals y F. C. Morabito (Eds.), *Recent advances in nonlinear speech processing* (Vol. 48, pp. 53-60). Cham: Springer. doi.org/10.1007/978-3-319-28109-4_6
- Carmona-Duarte, C., Plamondon, R., Gómez-Vilda, P., Ferrer, M. A., Alonso, J. B., y Londral, A. R. M. (2016). Application of the lognormal model to the vocal tract movement to detect neurological diseases in voice. En Y. Chen, S. Tanaka, R. Howlett y L. Jain (Eds.), *Innovation in medicine and healthcare* (Vol. 60, pp. 25-35). Cham: Springer. doi.org/10.1007/978-3-319-39687-3_3
- Carmona-Duarte, C., Gómez-Vilda, P., Plamondon, R., Rodellar-Biarge, V., Palacio-Alonso, D., Ferrer-Ballester, M. A., y Alonso-Hernández, J. B. (2017). Phonatory and articulatory correlates in kinematic neuromotor degeneration. En J. Ibáñez, J. González-Vargas, J. M. Azorín, M. Akay y J. L. Pons (Eds.), *Converging clinical and engineering research on neurorhabilitation II* (pp. 203-208). Cham: Springer. doi.org/10.1007/978-3-319-46669-9_36
- Delgado, C. (2016). Nadie está preparado para el diagnóstico de una enfermedad como la ELA. *Diario El País*. Disponible en http://economia.elpais.com/economia/2016/11/05/actualidad/1478300583_541697.html
- Gomez-Barrero, M., Galbally, J., Fierrez, J., Ortega-García, J., y Plamondon, R. (2013). Variations of handwritten signatures with time: A Sigma-Lognormal analysis. In *International Conference on Biometrics (ICB)*. doi.org/10.1109/ICB.2013.6612974
- Gómez-Vilda, P., Londral, A. R. M., Rodellar-Biarge, V., Ferrández-Vicente, J. M., y de Carvalho, M. (2015). Monitoring amyotrophic lateral sclerosis by biomechanical modeling of speech production. *Neurocomputing*, 151(P1), 130-138. doi.org/10.1016/j.neucom.2014.07.074
- O'Reilly, C., y Plamondon, R. (2009). Development of a Sigma-Lognormal representation for on-line signatures. *Pattern Recognition*, 42(12), 3324-3337. doi.org/10.1016/j.patcog.2008.10.017
- O'Reilly, C., y Plamondon, R. (2012). Design of a Neuromuscular Disorders Diagnostic System using human movement analysis. *The 11th International Conference on Information Sciences, Signal Processing and Their Applications*, 787-792.
- Peterson, G. E., y Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 24(2), 175-184.
- Plamondon, R. (1995). A kinematic theory of rapid human movements: Part I. Movement representation and generation. *Biological Cybernetics*, 72(4), 295-307. Disponible en <https://www.ncbi.nlm.nih.gov/pubmed/7748959>
- Plamondon, R., y Djioua, M. (2006). A multi-level representation paradigm for handwriting stroke generation. *Human Movement Science*, 25(4-5), 586-607. doi.org/10.1016/j.humov.2006.07.004
- Plamondon, R., Djioua, M., y Mathieu, P. A. (2013). Time-dependence between upper arm muscles activity during rapid movements: Observation of the proportional effects predicted by the kinematic theory. *Human Movement Science*, 32(5), 1026-1039. doi.org/10.1016/j.humov.2012.07.006
- Plamondon, R., O'Reilly, C., y Ouellet-Plamondon, C. (2014). Strokes against stroke - Strokes for strides. *Pattern Recognition*, 47(3), 929-944. doi.org/10.1016/j.patcog.2013.05.004
- Plamondon, R., O'Reilly, C., Rémi, C., y Duval, T. (2013). The lognormal handwriter: Learning, performing, and declining. *Frontiers in Psychology*, 4(DEC), 1-14. doi.org/10.3389/fpsyg.2013.00945